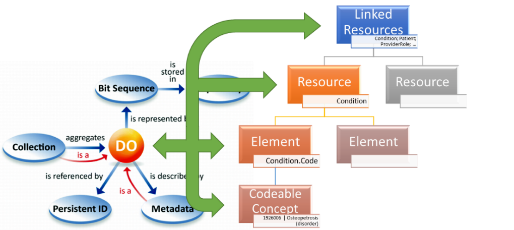
### The Problem

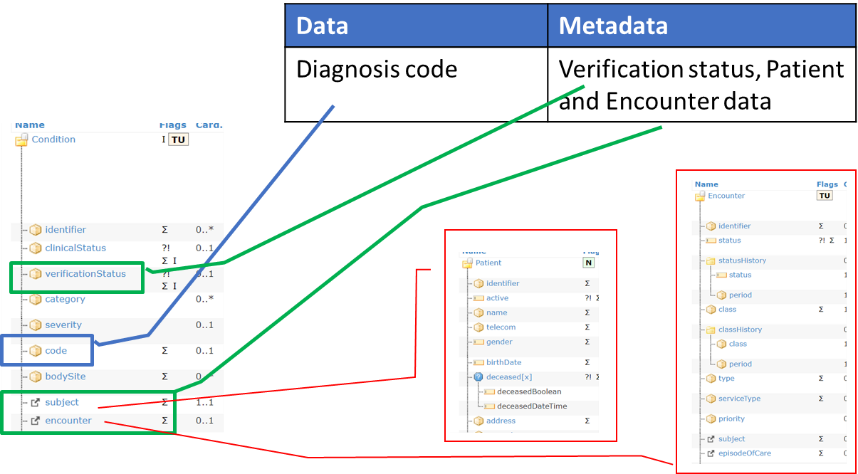
The concept of FAIR digital object (FAIR DO) is quite wide and can cover from a single atomic information (e.g., a coded diagnosis) up to a collection of data (e.g., a data set). Data, moreover, can represent quite different kinds of information: it might be for example a waveform, an image, a condition, a medication, or other kinds of data.

This variety makes not so straightforward the mapping with the FHIR resources covering a large range of resources and elements (see figure below).



**Figure 1 – FAIR Digital Object representation**

Similar considerations can be done also for the metadata, where, depending on what the metadata is, the metadata elements can be represented within the same resource documenting the data, or by a set of linked resources.

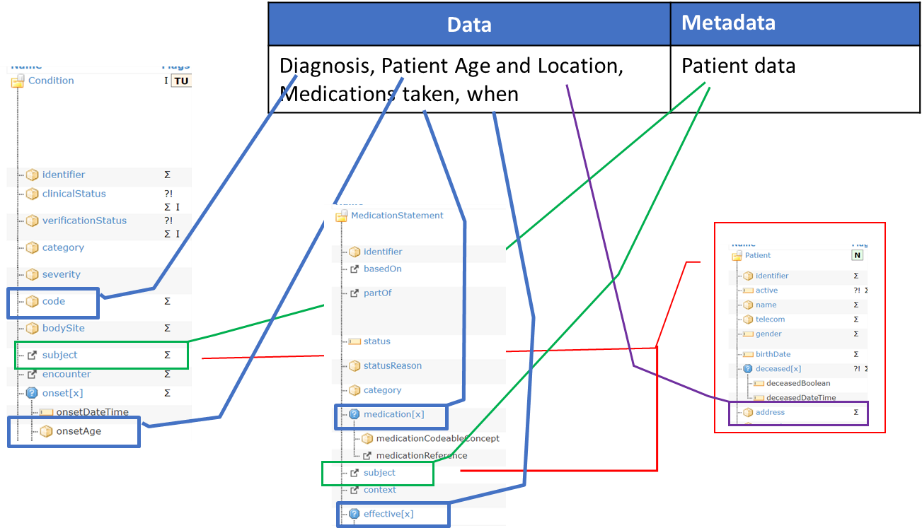


**Figure 2 – Mapping FAIR data and metadata**

#### Data vs metadata

The logical distinction between metadata and data (both are FAIR DO) is a key point; with metadata playing a key role for making data findable, accessible, interoperable and reusable. Metadata are usually described as the “data about data”, but in practice what is “data” and what is “metadata” is a matter of perspective: based on the context, the same information can be considered as part of the data or of the metadata. This makes even more complex the mapping to the FHIR resources.

The figure below shows how information that were part of the metadata in the previous example are now part of the data; and how metadata and data elements are represented in the same FHIR resource.



**Figure 3 – Mapping FAIR data and metadata**

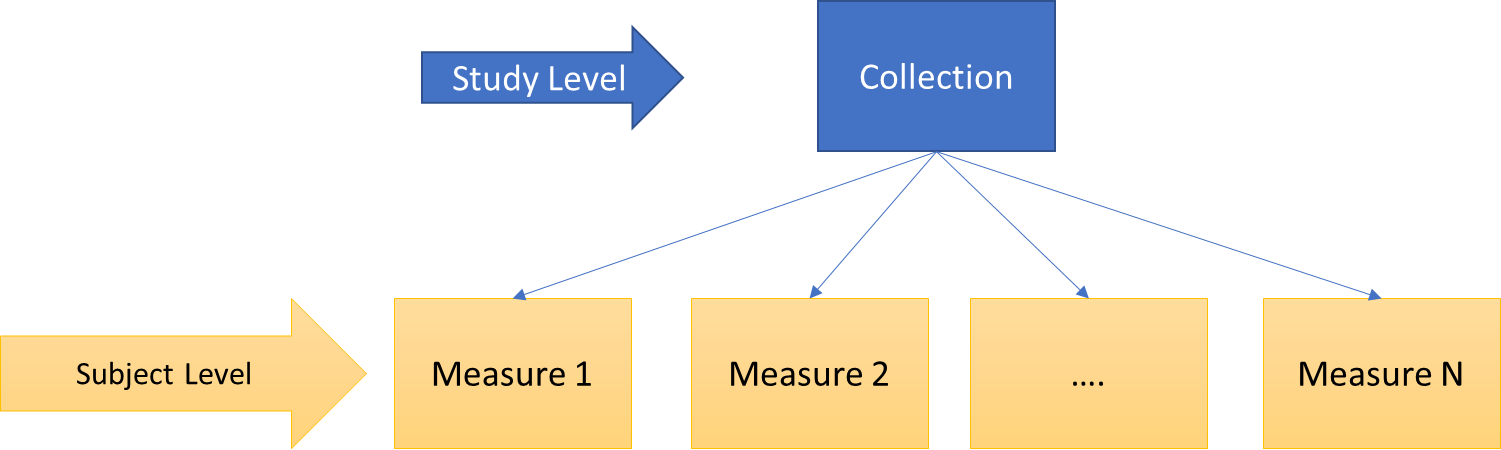
#### Metadata identification

In consideration of the previous arguments, the expectation of having a persistent and unique ID that identifies the metadata FAIR data object, distinct from the data FAIR data object, cannot be always satisfied in the FHIR space, except for specific contexts.

### The approach

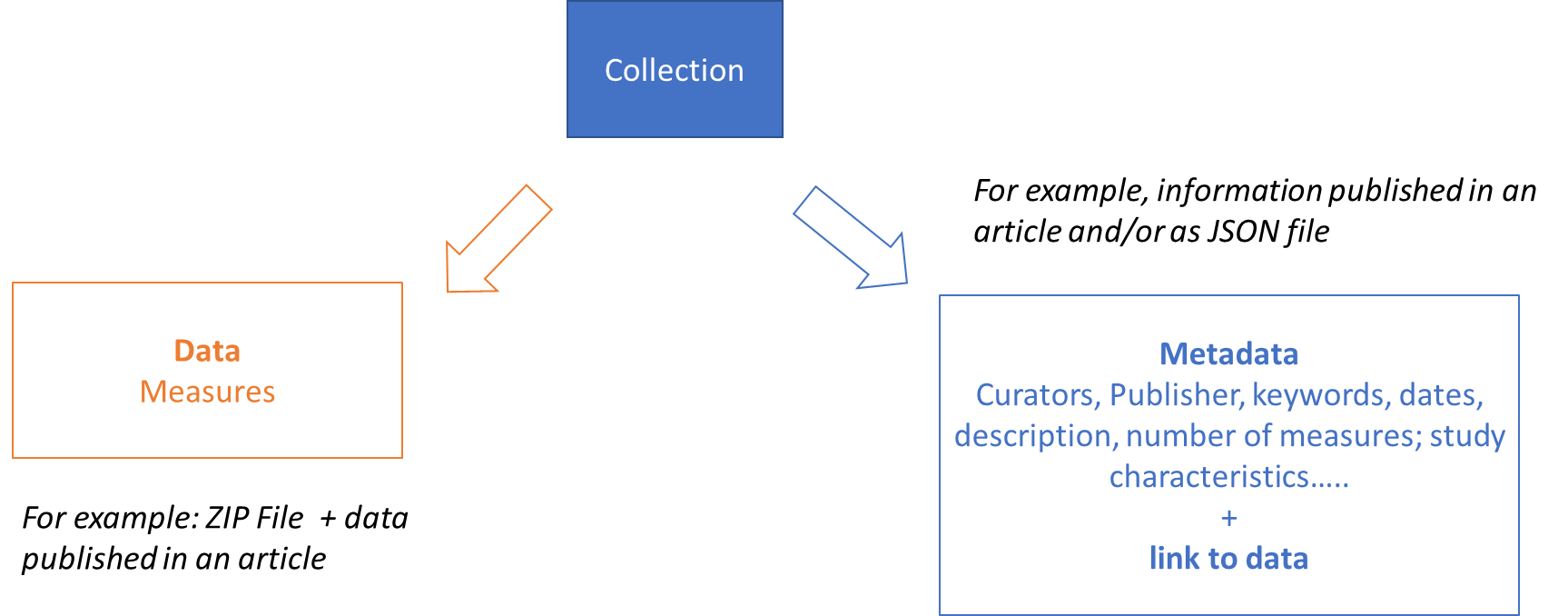
Even though many levels of granularity can be identified, in this work we consider two main levels:

1. The **study level**: describing the collection of data referring to a specific study, publication, usage context.
2. The **subject Level**: providing the metadata/data information associated to a single subject.



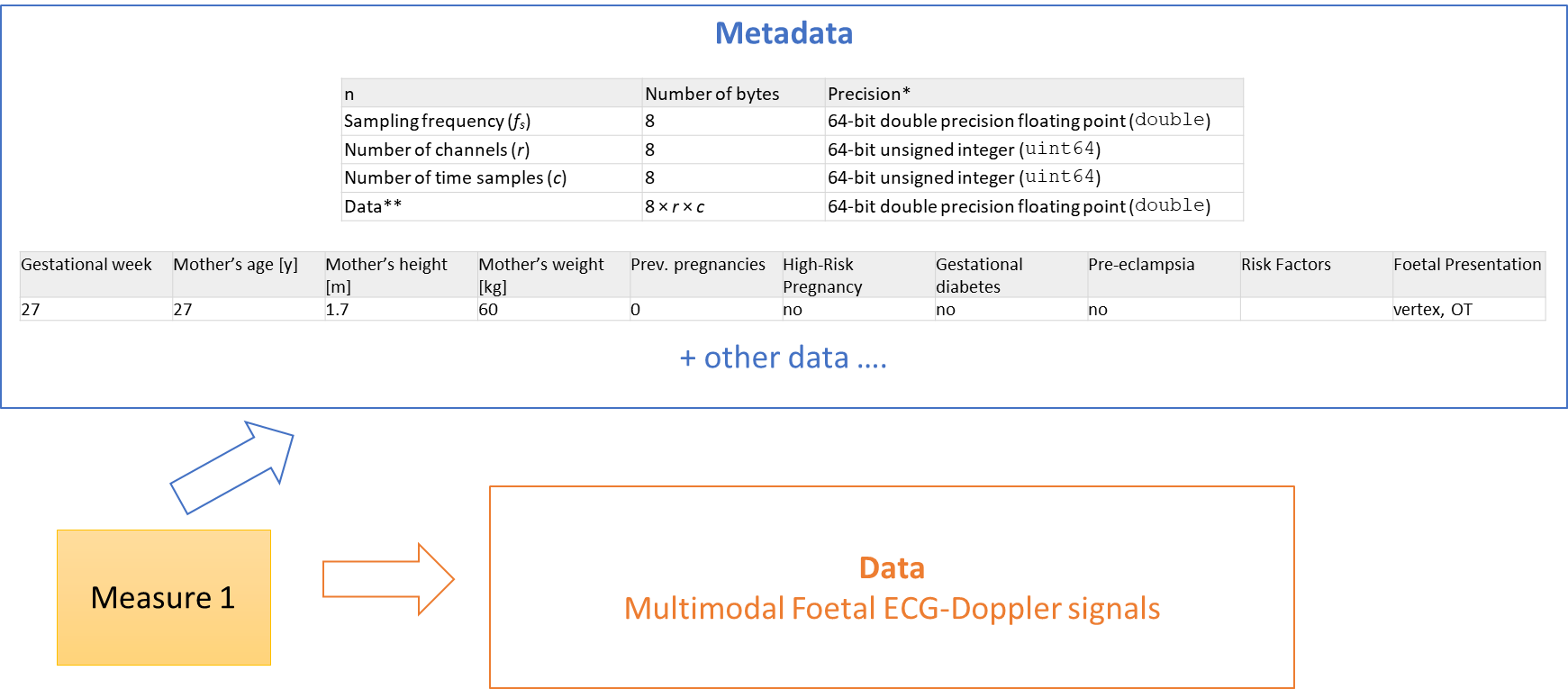
**Figure 4 - Level of data objects granularity considered in this guide**

For collections of data (study level) it is reasonable to assume that a distinction between metadata and data is possible.



**Figure 5 – Metadata and data for a data collection.**

This might be not so true for the instance level, where information as the gestational week or the mother height and weight in the example below, might be considered metadata for the signals, but also patient level data.



**Figure 6 – Metadata and data at the subject level.**

### Metadata Representation in FHIR

#### Study Level (Collection)

HL7 FHIR provides several candidates resources to represent metadata of collection of data (e.g., Bundle, Lists..) and the choice of the resource to be used strongly depend on the usage context. **Library,** **Citation** and **ResearchStudy**, seem to be the most promising resources to better support the “rich metadata” FAIR requirement;however not all of them are available in the current FHIR Version. Furthermore, due to their low FHIR maturity level, they may be subject to non-negligeable changes among the different FHIR versions, including the way the linkage between metadata and data is realized.

The kinds of data objects that the metadata may refer to can be:

1. Non-FHIR objects
2. FHIR and non-FHIR objects
3. FHIR-only objects

Depending on the “metadata” resource, the FHIR version chosen and the type of object to point to, different solutions can be adopted. For example, the current Library resource covers this requirement through the content element (Attachment datatype), in this case the link to data is recorded by using an uri (this may or may not be enough for referring FHIR resources). Citation (looking at the FHIR R5 specs) provides instead more flexible and comprehensive solutions through the citedArtifact.relatedTo.target[x] element; providing different means (uri, resource references, business identifiers) to handle these references.

It’s worth to mention that FHIR specifies a dedicated resources (Provenance) to record Provenance information typically part of the “metadata”.

#### Subject Level (e.g. Patient)

As described above, at the subject level, the boundary between metadata and data is not always so sharp.

This guide will not therefore attempt to **prescribe any tight separation of metadata/data at the subject level** when FAIR is implemented by using HL7 FHIR.

There are HL7 FHIR resources (e.g. EvidenceReport, Bundle[collection], Composition, List; DocumentManifest) that can be potentially used to record some subject level metadata information as a distinct FHIR resource instance. Nevertheless, typically metadata information is recorded by a set of linked FHIR resources. For example, a Condition instance documenting a specific problem (the data), may use the *subject*, the *encounter* and the *evidence* references to Patient, Encounter and Observation FHIR instances to document the patient, encounter and other information used as evidence to describe the context of this problem (the metadata).

In case of *non-FHIR data objects* or mixed cases (*FHIR and non-FHIR objects*) implementers should evaluate the feasibility and the cost/benefit of transforming these objects (completely or partially) in FHIR instances; and/or creating FHIR resources (e.g. DocumentManifest) to document some metadata information.